Assessing the Availability of Crowdsourced Street-Level Imagery for Urban Perception Mapping

Ayse Giz Gulnerman(✉)

Ankara Hacı Bayram Veli University, Ankara 06560, Türkiye
ayse.gulnerman@hbv.edu.tr

Abstract. The integration of urban science with spatial artificial intelligence, referred to as Urban-AI or Geo-AI, is playing an increasingly important role in monitoring cities, supporting urban planning, and addressing urban challenges. In this context, investigating how cities are perceived by individuals and understanding the effects of urban environments on people have emerged as important research topics that can directly influence individuals' quality of life. In recent years, street-level imagery (SLI) has been widely used to evaluate urban perception through artificial intelligence-based algorithms. In particular, generative artificial intelligence offers growing potential for analyzing urban environments and supporting the development of spatial solutions. However, ethical risks such as algorithmic bias, socio-economic inequalities, and data representativeness issues remain important concerns. Therefore, before enabling generative AI-based systems to support transformations in urban environments, it is critical to investigate their functionality and development from a spatial data perspective. This study presents an empirical framework to evaluate the usability of crowdsourced SLI for urban perception research.

Keywords: Street-Level Imagery (SLI) · Urban Perception Mapping · Geo-AI · Crowdsourced Data · Data Representativeness · Spatial Data Quality

1 Introduction

With the intensification of global urbanization, challenges related to urban form, function, and sustainability are becoming increasingly complex and urgent [1, 2]. These interactions between urban science and artificial intelligence, referred to as UrbanAI or Geo-AI, enable high-resolution and dynamic monitoring and modeling of urban processes [2]. These emerging approaches highlight that traditional physical planning systems were not designed to address the crises of today's complex world and therefore require radical transformation [3]. Data-driven generative artificial intelligence (Generative AI) and large language models (LLMs) provide new opportunities not only to explain existing urban systems but also to invent new urban forms that optimize multiple criteria [4]. However, serious ethical concerns accompany these technological developments, including algorithmic bias, the deepening of social inequalities, and potential human

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16765, pp. 395–408, 2027.
https://doi.org/10.1007/978-3-032-30536-7_25

rights violations [1, 5]. In particular, the training of urban AI systems on datasets reflecting "WEIRD" (Western, Educated, Industrialized, Rich, Democratic) values increases the risk of discrimination and exclusion on a global scale [5]. For the future of cities, an interdisciplinary approach that synthesizes urban science, computer science, and geographic information science (GIScience) is proposed [1]. This approach advocates balancing technological innovation with principles of social justice, sustainability, and human-centered design, ensuring that smart interventions are grounded in the real needs and experiences of urban residents [1, 6]. Consequently, it is essential to establish a comprehensive research framework that aligns emerging digital tools with societal needs and to structure these technologies within accountable governance mechanisms [1, 2, 6].

In the 21st century, the massive data generated by the widespread use of smartphones and applications together with satellite imagery, IoT sensors, social media, and Street-Level Imagery are regarded as major data sources for urban analytics [1, 2, 6]. Based on the idea that cities are computationally measurable, artificial intelligence technologies are widely recognized as important tools for making sense of data derived from these heterogeneous sources [1, 6].

Street View Imagery (SVI) typically refers to platform-specific, often panoramic and vehicle-based images such as those from Google Street View, whereas Street-Level Imagery (SLI) is a broader concept that includes images captured at street level from various sources and devices (e.g., Mapillary, KartaView, mobile cameras). In this study, the term SLI is preferred in sections beyond the literature review, as it provides a more comprehensive and inclusive framework.

SLI has become firmly established as an integral element of urban analytics and GIScience, with advances in artificial intelligence enabling the automation of geo-information extraction and increasing the spatial coverage provided by data sources [7]. However, previous studies indicate that certain regions are disadvantaged in SLI based research due to limitations in spatial and temporal coverage. In particular, developing regions, rural areas, and parts of South America and Africa remain underrepresented, highlighting geographic inequalities in data availability and research focus [7]. Biljecki and Ito (2021) show that 80% of studies using SLI focus on only a single city. In addition, numerous SLI data sources exist; however, the majority offer spatial coverage limited to one or a small number of countries [7].

1.1 Studies in Türkiye

Building on the general literature, this study focuses on the Turkish context by presenting methods to assess the availability and usability of street-level imagery (SLI) in Türkiye, while also reviewing the existing body of SLI-related studies conducted in the country.

Kassab and Uzun (2024) examine urban segregation between two districts in Istanbul by combining semantic segmentation of street view images with perceptual scene ratings [8].

Gulnerman (2025) evaluates the ability of generative AI models (Gemini 1.5 Flash and LLaVA 13B) to assess urban perception from street-level images in a small central region of Ankara and finds that Gemini provides more consistent multilingual results, highlighting the potential of GenAI for automated urban perception mapping [9].

Küçükyağcı (2025) investigates changes in the green view index (GVI) at street level in a street in Kadıköy, Istanbul, using Google and Yandex street-level imagery, demonstrating that the green texture decreased due to urban transformation [10].

Iban et al. (2025) argue that available benchmark datasets for urban semantic segmentation are biased, as they are largely based on European and North American urban patterns (e.g., Cityscapes) and are primarily designed for navigation or driving purposes (e.g., KITTI, CamVid, BDD100K), limiting their applicability to diverse urban contexts. To address this issue, they introduce a new SVI dataset from Mersin that reflects diverse urban typologies in Türkiye [11].

Gönül and Durak (2025) evaluate human perception along a high street in Bursa adjacent to a historical area listed as a UNESCO heritage site, assessing changes following urban transformation between 2000 and 2022 using Google Street View (GSV) imagery from 2014, 2018, 2020, and 2023 [12].

Most studies conducted in the Turkish context focus on a limited part of a city or only a few streets. Among these, only the study by Iban et al. [11] adopts a comprehensive data collection approach, aiming to develop a semantic segmentation benchmark better aligned with Türkiye's urban morphology. The existing literature indicates that street-level imagery (SLI) research in Türkiye is still a relatively new and emerging field. Therefore, providing a preliminary assessment of data availability and quality in Türkiye is considered essential for guiding future research in this area.

1.2 SLI Assessment Studies in Literature

Studies evaluating SLI are systematically presented in Table 1 with respect to the data sources employed, the spatial scope of the analyses, and the methodological approaches adopted.

SLI is typically collected from a moving object (e.g., a motor vehicle, bicycle, or pedestrian), capturing images along the movement trajectory at a consistent angle (or panoramically) [19]. The resulting data has been evaluated in many studies in terms of spatial coverage [14, 16, 17]. This coverage is commonly assessed by examining whether images are present within specific road segments, grid cells, or point-based buffer zones. However, the presence and diversity of objects within the imagery have largely been overlooked in the literature. While the high availability of images along major roads provides valuable contributions to applications such as road perception for motorized travelers and the development of autonomous driving algorithms, these images alone are not sufficient to represent the urban perception of residents living in the area. Therefore, the content of the images and the capture perspective should be evaluated in accordance with the research theme, and a content-oriented perspective should be incorporated into coverage analyses.

1.3 Aim of the Study

This study aims to evaluate the usability, availability, and quality of crowdsourced SLI for urban perception research in Türkiye by proposing a methodological framework and providing a preliminary assessment to support future Geo-AI and urban analytics

Table 1. Street view imagery data assessment studies.

Author	Used SLI source	Geospatial scope	Methodology
Juhász and Hochmair [13]	Mapillary and GSV	Global evaluation for main roads, detailed analysis for 11 urban, rural and mixed areas in USA and Europe for various road types	User contribution pattern and spatial coverage analysis based on tile.
Quinn and León [14]	Mapillary, GSV, OpenStreetCam	24 major world cities and 25 differently sized cities in Brazil	Coverage analysis (qualitatively assessed by three analysts), along with comparisons between platforms and an examination of the reasons behind missing data regions.
Ma et al. [15]	Mapillary	Worldwide dataset covering 5 years (2014–2019)	Explorative analysis-User sequence-based contribution activity analyzed based on different photo collection modes (driving, walking, cycling)
Fry et al. [16]	GSV	371 cities in Latin America	Image availability, Image age, age variability (standard deviation) over points sampled 500 m intervals
Mahabir et al. [17]	Mapillary, OpenStreetCam	4 cities in USA	Coverage analysis with 1kmx1km grid (cumulative sequence based)
Biljecki and Ito [7]	>600 academic papers	80% of papers focus solely on one city	Review academic paper

(*continued*)

applications, which will be further supported by an exploratory analysis based on SLI data from Ankara.

Table 1. (*continued*)

Author	Used SLI source	Geospatial scope	Methodology
Hou and Biljecki [18]	Mapillary	The Kowloon area (Hong Kong) for a comprehensive evaluation, along with eight additional cities worldwide to assess selected quality elements (e.g., Singapore, Chicago, Santiago, Zagreb)	7 quality categories (including image, spatial, and temporal quality), comprising 48 quality elements.
Fan, Feng, Biljecki [19]	GSV	62% of buildings in London	Element-level SVI coverage estimation

2 Methodology

The methodological framework is structured around four main components: Spatial Distribution and Coverage Assessment, Temporal Data Assessment, Participation-Based Assessment, and Domain (Theme)-Based Assessment.

2.1 Spatial Distribution and Coverage Assessment

Spatial coverage is addressed within the context of the road network and incorporated into the analysis. OpenStreetMap (OSM) road data for the study area are retrieved using the OSMnx library and transformed into an analysis-ready format. The road network is divided into 100-m segments to establish comparable and homogeneous spatial units. Each SLI point is associated with its nearest road segment using a nearest neighbour approach, with the linkage defined based on a distance threshold ($\leq$100 m).

While no standard segment length is adopted in the literature, previous studies on street-level imagery commonly rely on spatial aggregation and proximity-based association methods to derive meaningful analytical units [17–19]. In this study, a 100 m segmentation was selected as a pragmatic balance between spatial resolution and analytical robustness.

Within this framework, the number of SLI points per road segment is calculated to define street-level coverage. The resulting values are classified using the K-means clustering method to identify different levels of data density. Segments with no associated data are retained within the analysis, enabling the identification and evaluation of spatial gaps in data coverage.

2.2 Temporal Data Assessment

Temporal characteristics of the SLI data are examined using metadata-derived timestamp information. The timestamp attribute, stored in milliseconds, is retrieved from

the database and converted into a standard datetime format to enable temporal analysis. Records with invalid or missing timestamps are excluded to ensure data consistency. Following this transformation, temporal attributes are extracted at the year level, allowing the distribution of SLI data over time to be systematically evaluated.

The number of images is aggregated for each year to identify temporal trends, data accumulation patterns, and potential gaps in data availability. This approach enables the assessment of data recency and temporal representativeness, providing insight into whether the dataset reflects current urban conditions or is dominated by imagery from specific time periods.

2.3 Participation-Based Assessment

To analyze the distribution of participation, the number of images produced per user is calculated and ranked according to contribution levels. This approach reveals whether data production is concentrated among a limited number of users. In addition, the representativeness of the data production structure is evaluated based on the most active contributors.

To assess the level of inequality in contribution, cumulative contribution shares of users are calculated and evaluated using a Pareto-based approach. This analysis identifies the extent to which data production is dominated by a small number of contributors and provides insight into the balance of participation.

To examine temporal participation dynamics, user contributions are disaggregated by year, and the yearly distribution of contributions by leading users is analyzed. This enables the evaluation of temporal continuity in data production, user-based contribution patterns, and potential periods of intensified activity.

This comprehensive approach aims to reveal not only the spatial characteristics of SLI data but also its participation-based production structure and representational equity.

2.4 Domain (Theme)-Based Assessment

Domain-based assessment is grounded in the analysis of the thematic content of SLI data through the objects and visual elements. In this context, image content is evaluated using computer vision-based methods, particularly semantic segmentation and object detection techniques.

Through semantic segmentation, key components of the urban environment (e.g., greenery, buildings, roads, sidewalks, sky, vehicles, and pedestrians) are identified at the pixel level. The proportional distribution of these classes is calculated for each image, enabling the quantification of thematic representation across the dataset.

Object detection is employed to identify the presence and frequency of specific urban elements. This approach provides complementary information on scene composition, allowing for a more comprehensive evaluation of thematic content. The resulting thematic indicators are associated with spatial analysis units such as road segments or grid cells. This enables the examination of spatial variations in thematic distributions as well as potential biases in data representation.

This methodological approach aims to reveal the extent to which SLI data represent the visual and functional characteristics of the urban environment and to assess its applicability for urban perception and Geo-AI-based analyses.

Mapillary provides semantic segmentation outputs for SLIs. Through the Mapillary API, geometric and descriptive information about objects present in the images was obtained. Using the image width and height, the pixel area of each detected object was calculated and expressed as a percentage of the total image area. These values were stored both at the detection level and as image-level summaries based on object classes. In this way, each image was represented not only by the presence of objects but also by the proportional area occupied by each object category. This structure formed the basis for deriving thematic components such as road, sidewalk, building, vegetation, sky, vehicle, and person, which were subsequently used in the calculation of the usability score.

The usability score (US) is based on three main indicators: urban visibility (UV), obstruction (OB), and diversity (DV). Urban visibility represents the proportion of key perceptual elements in the image; obstruction reflects the extent to which elements such as vehicles and people block the view; and diversity represents the number of different object groups present in the image. The proposed usability score does not directly follow an existing standardized metric in the literature. Instead, it was developed as a heuristic composite index based on semantic segmentation outputs and informed by previous studies on street view imagery quality, urban perception, and content-based evaluation frameworks (Hou & Biljecki, 2022; Ito & Biljecki, 2021).

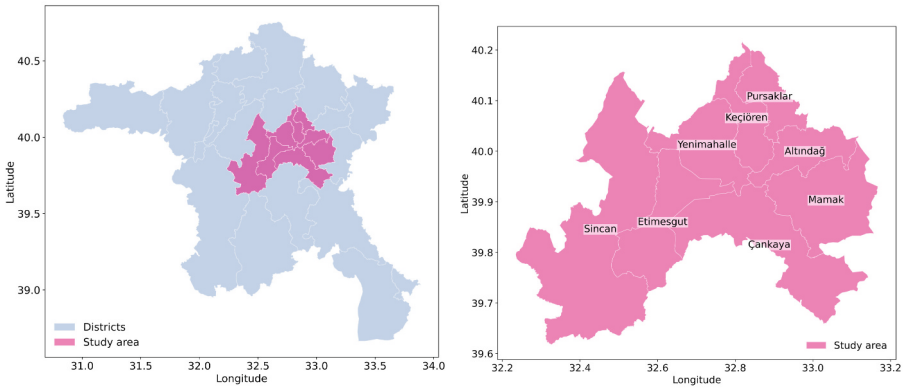
3 Case Study and Results

In this study, the city of Ankara was selected as the study area. Ankara is not only the capital of Türkiye but also the second most populous city in the country. Although many studies in the literature assessing SLI quality have focused on major cities worldwide, Ankara has not been included in these analyses. In this context, examining Ankara is considered to partially represent cities in Türkiye in terms of SVI quality.

Figure 1 illustrates all districts of Ankara and the spatial location of the study area. The study area is selected from districts characterized by an urban population density exceeding 500 persons per km^2, ensuring the representation of relatively dense urban environments.

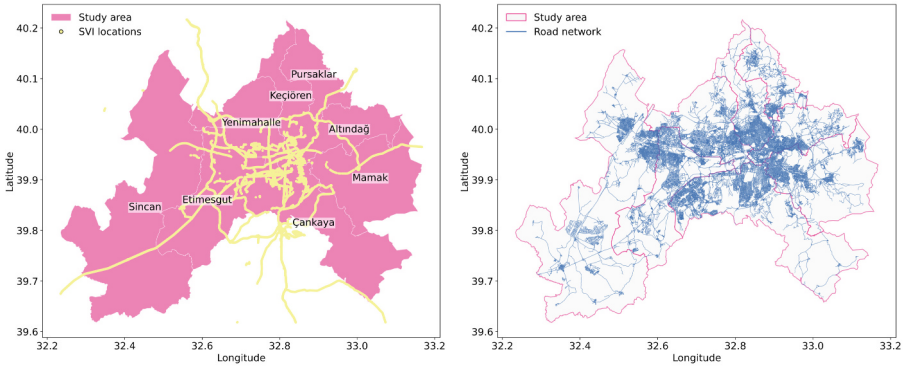
Mapillary SLI data were retrieved using the Mapillary API within the bounding box of the study area and subsequently refined through a spatial intersection with the study area boundaries. The resulting dataset is presented in Fig. 2(a). In addition, the road network within the study area was obtained from the OpenStreetMap (OSM) platform and visualized in Fig. 2(b). A partial spatial correspondence can be observed between these two datasets, particularly along primary road segments where SLI data are more consistently available. In contrast, the presence of SLI data along secondary roads appears to be limited and is mainly concentrated in areas closer to the urban core.

The spatial distribution of SLI data across the road network, classified using the K-means algorithm, indicates a highly uneven coverage structure (Fig. 3). A substantial portion of the road network contains no SLI observations, highlighting significant gaps in data availability. Primary arterial roads predominantly fall within low to moderate



(a) (b)

Fig. 1. (a) Administrative boundaries of all districts in Ankara; (b) districts selected as the study area.



(a) (b)

Fig. 2. Spatial distribution of (a) SLI locations and (b) road network across the area.

density classes (e.g., 1–16 and 16–50 images per 100 m segment), forming the main backbone of SLI data. In contrast, higher-density classes do not exhibit spatial continuity and are largely confined to short road segments or point-like clusters, particularly in areas close to the urban core. These findings suggest that while major transportation corridors are relatively well represented, the overall spatial distribution of SLI data remains fragmented and uneven. High-density observations tend to occur as localized concentrations rather than along continuous street networks, reflecting the opportunistic nature of data collection and the lack of uniform spatial coverage.

The temporal distribution (Fig. 4) shows a sharp peak in 2019, indicating a major data collection period, followed by a decline and a gradual increase in recent years. Overall, the dataset is temporally uneven, with a strong concentration in a single year.

Figures 5, 6, and 7 collectively reveal that SLI data production is highly concentrated among a limited number of contributors and exhibits notable temporal variability.

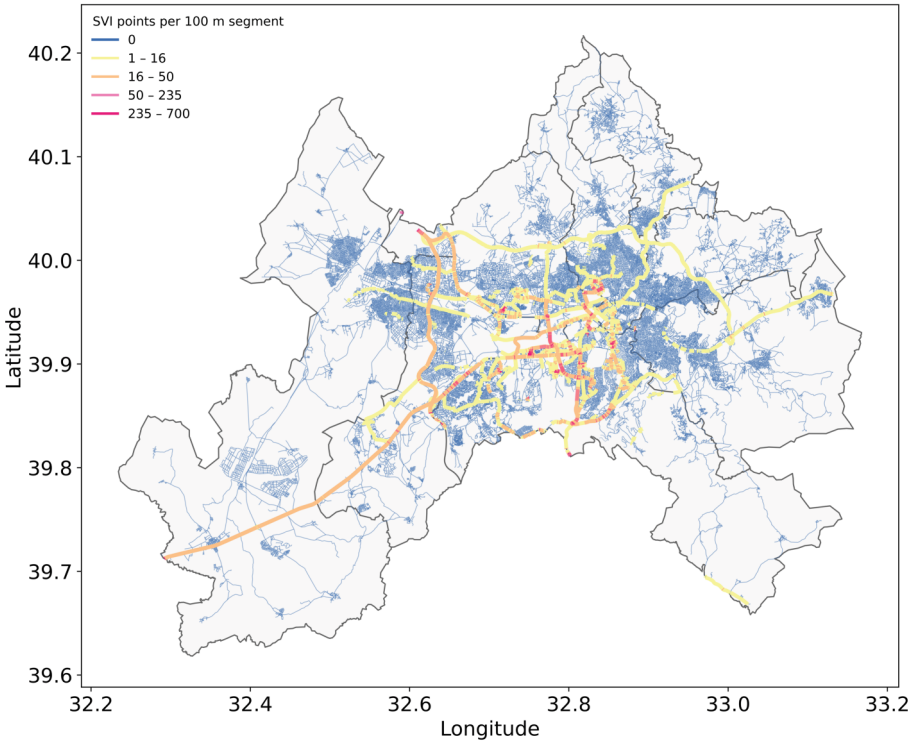


Fig. 3. SLI street-based coverage map.

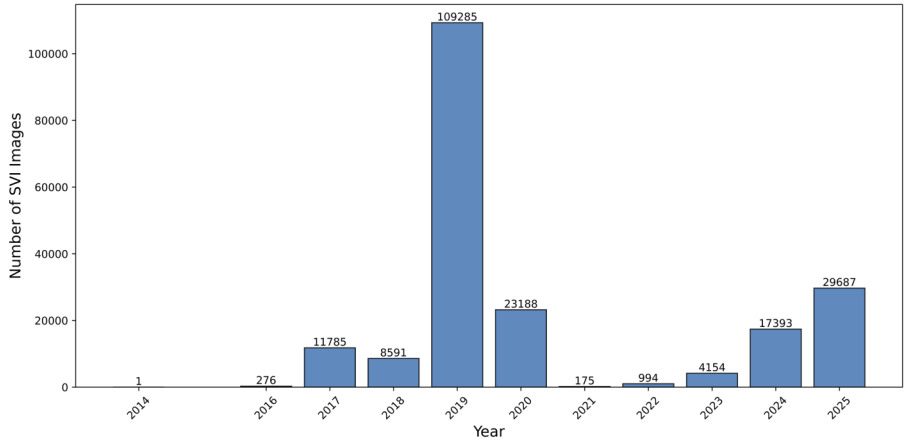


Fig. 4. Temporal distributions of SLI data by year.

Figure 5 shows that a small group of users produces a disproportionately large share of the total images, indicating an uneven participation structure. This pattern is further

confirmed by the cumulative contribution curve in Fig. 6, where a relatively small number of contributors accounts for the majority of the dataset, reflecting a clear Pareto-like distribution.

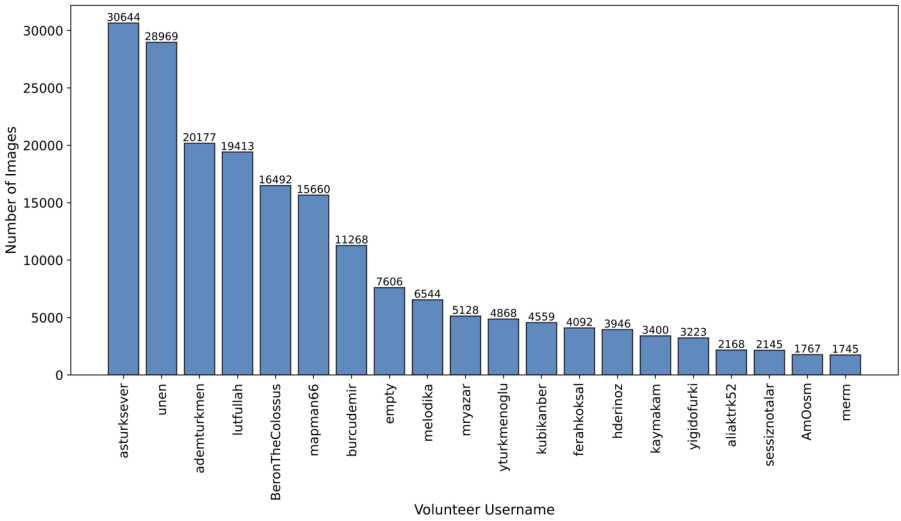


Fig. 5. Top Contributors by Number of SLI Images

Figure 7 adds a temporal dimension to this pattern, demonstrating that leading contributors dominate data production across multiple years, although their activity levels vary over time. This suggests that SLI data generation is not only user-dependent but also temporally inconsistent, relying on intermittent contributions from a core group of active users. Overall, these findings highlight that SLI data production is both participant-driven and unevenly distributed, raising important considerations regarding data representativeness and potential biases in urban analyses.

An object-based framework to evaluate the suitability of street-level imagery for urban perception analysis is applied. Using Mapillary data, detected objects are first grouped into perception-relevant categories such as road, building, vegetation, sky, vehicle, and pedestrian elements. For each image, the spatial composition of these categories is quantified based on their area coverage percentages. Building on this composition, a usability score is constructed using three main components: (i) urban visibility, representing the presence of key environmental elements such as roads, buildings, vegetation, and sky; (ii) diversity, reflecting the variety of perceptual cues within the scene; and (iii) obstruction, penalizing elements that reduce scene readability, such as vehicles and pedestrians. In addition, images with insufficient informative content are further penalized. The resulting score is normalized to a 0–100 range, and images are classified into different usability levels, enabling the transformation of raw street-level imagery into a dataset suitable for urban perception analysis.

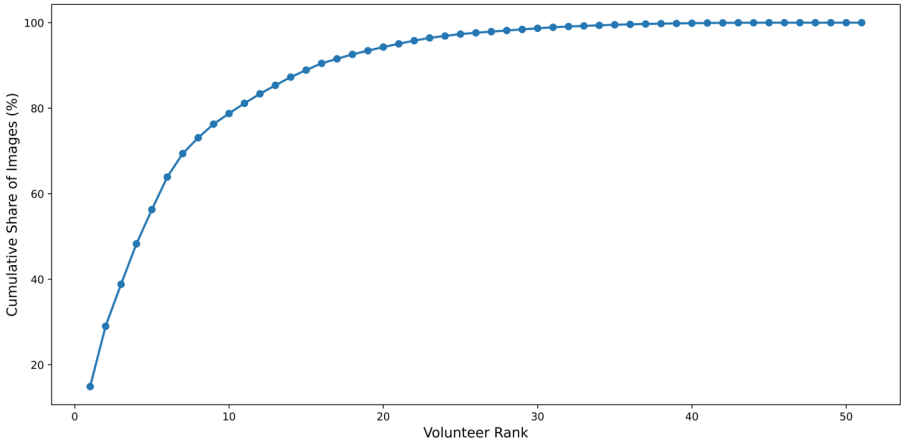


Fig. 6. Cumulative Contribution of Creators.

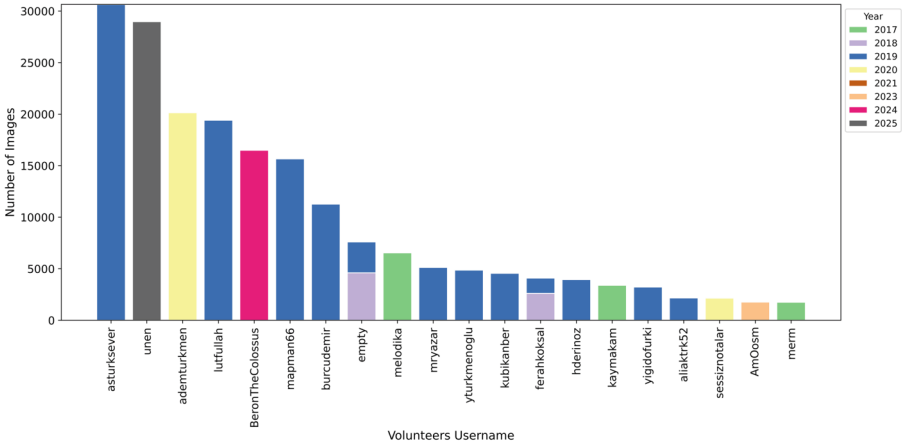


Fig. 7. Yearly Contributions of Top Volunteers.

The distribution indicates a strong concentration of images in the medium usability range, while a substantial proportion of images exhibit very low usability scores, highlighting the presence of noisy or non-informative observations in the dataset (Fig. 8(a)). The scatter plot (Fig. 8(b)) further demonstrates a strong negative relationship between obstruction and usability, confirming that visual occlusions significantly reduce the analytical value of SLI. However, the variability observed at low obstruction levels suggests that usability is not solely determined by occlusion, but also by the overall spatial composition of the scene.

The correlation analysis reveals that usability is primarily driven by scene openness and structural visibility, with sky and road coverage emerging as the strongest predictors (Fig. 9). Vegetation and vehicle presence also contribute positively, suggesting that both environmental quality and urban activity enhance scene interpretability. In contrast,

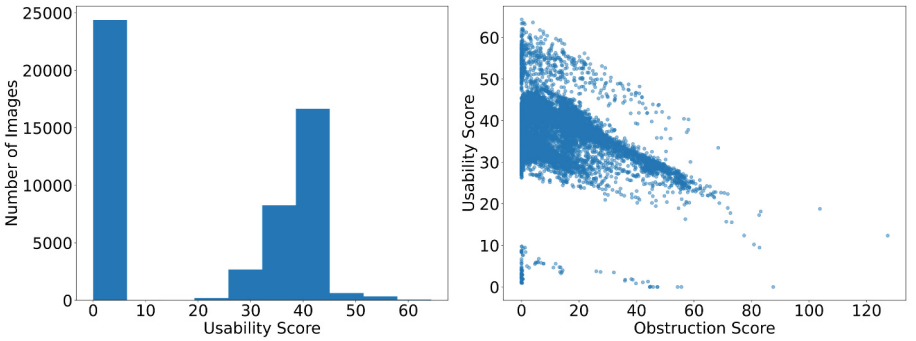


Fig. 8. SLI content based distribution of (a) usability score (b) obstruction on usability.

pedestrian presence shows minimal influence, indicating that usability is more dependent on spatial composition than human activity density.

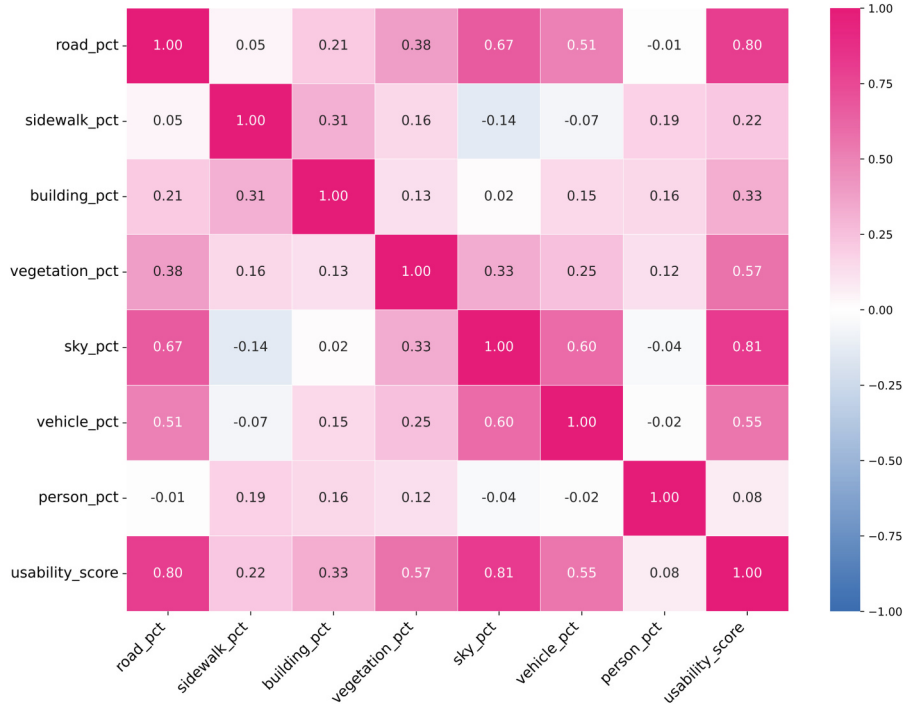


Fig. 9. Correlation between object coverage and usability.

4 Discussion and Conclusion

This study explores the adequacy of SLI-based data for urban perception mapping through a series of exploratory analyses. Districts of Ankara characterized by high urban population density are examined in terms of data availability within Mapillary, a crowdsourced SLI platform. As one of the first studies in Türkiye to assess SLI data quality, the findings indicate that data availability is largely concentrated along primary arterial roads, while street-level or point-based data is mainly observed in areas that can be considered part of the urban core.

From a temporal perspective, the dataset exhibits significant variability, with certain years showing peak levels of data production while others contain little to no data. In addition, data production is highly uneven in terms of participation, with approximately 20% of users contributing the majority of the dataset. These findings highlight notable temporal and participation-based biases in the data.

Within the scope of thematic assessment, the visual content of the images is analyzed by quantifying the area occupied by different objects, enabling the derivation of usability scores. Furthermore, an obstruction score is introduced to evaluate the extent to which dominant elements within the images may hinder meaningful interpretation.

Overall, the analysis of approximately 200,000 images for Ankara suggests that, from a spatial perspective, the dataset does not adequately represent a large portion of the urban environment. The data is also characterized by temporal and participation-related biases, and a substantial proportion of the images exhibit limited suitability for urban perception studies.

Nevertheless, the presence of repeated observations in central urban areas and relatively high usability scores in certain high-density zones indicate that crowdsourced SLI data retains potential for targeted analyses. Despite limitations in data quality, biases, and content-related constraints, openly available SLI datasets can still serve as valuable resources for localized studies. Prior to their application, data filtering, weighting, integration, and enrichment techniques can be employed to enhance their usability and analytical value.

In future studies, defining the usability score through an explicit formulation and weighting scheme, and validating it using human-based evaluations and established urban perception models, is considered an important direction for improving methodological robustness.

Acknowledgments. The authors would like to thank Mapillary for providing open access to street-level imagery data and OpenStreetMap (OSM) for offering openly accessible road network data, both of which enabled the use of open science resources in this study. The authors also acknowledge the support of ChatGPT in proofreading the manuscript and assisting with Python coding throughout the research process.

This study was supported by the Scientific and Technological Research Council of Türkiye (TÜBİTAK) under the 2224-A - Grant Program for Participation in Scientific Meetings Abroad. The authors gratefully acknowledge this support.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Shi, W., et al.: The science of urban informatics. Sci. Bull. (2025)
2. Knoblauch, S., et al.: Urban AI for a sustainable built environment: Progress and future directions. Environment and Planning B: Urban Analytics and City Science 23998083261418565 (2026)
3. Batty, M.: Is physical planning over? If so, what will replace it? Environ. Plan. B: Urban Anal. City Sci. **52**(1), 3–6 (2025)
4. Batty, M.: Generative AI. Environ. Plan. B: Urban Anal. City Sci. **52**(5), 1031–1034 (2025)
5. Zhou, K., Constantinides, M., Quercia, D.: Should LLMs be WEIRD? Exploring WEIRDness and human rights in large language models. In: Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, vol. 8, p. 3 (2025)
6. Weng, Q., et al.: Urban visual-spatial intelligence: linking human and sensor perception for sustainable urban development. npj Urban Sustain. **5**(1), 65 (2025)
7. Biljecki, F., Ito, K.: Street view imagery in urban analytics and GIS: a review. Landsc. Urban Plan. **215**, 104217 (2021)
8. Kassab, J., Uzun, C.: Quantifying the Visuo-perceptual segregation between Ortabayır and Levent District in Istanbul through semantic segmentation. Sakarya Univ. J. Sci. **28**(6), 1165–1177 (2024)
9. Gulnerman, A.G.: AI-driven urban perception mapping: comparing Gemini and LLaVA. In: International Conference on Computational Science and its Applications. Springer Nature Switzerland, Cham (2025)
10. Küçükyağcı, P.Ö.: Examining the change of green view index (GVI) at street eye-level with GSV and YSV. GRID-Architect. Plan. Des. J. **8**(1), 263–282 (2025)
11. Iban, M.C., et al.: PalmCity: an emerging benchmark dataset for semantic segmentation of panoramic street view images in under-represented developing countries. In: International Conference on Computational Science and its Applications. Springer Nature Switzerland, Cham (2025)
12. Gönül, A., Durak, S.: On measuring the change in historical city centres: an attempt at comparing human perception and deep learning through visual quality of street space. Multimed. Tools Appl. **84**(32), 39283–39305 (2025)
13. Juhász, L., Hochmair, H.H.: User contribution patterns and completeness evaluation of Mapillary, a crowdsourced street level photo service. Trans. GIS. **20**(6), 925–947 (2016)
14. Quinn, S., León, L.A.: Every single street? Rethinking full coverage across street-level imagery platforms. Trans. GIS. **23**(6), 1251–1272 (2019)
15. Ma, D., et al.: The state of mapillary: an exploratory analysis. ISPRS Int. J. Geo Inf. **9**(1), 10 (2019)
16. Fry, D., et al.: Assessing Google street view image availability in Latin American cities. J. Urban Health. **97**(4), 552–560 (2020)
17. Mahabir, R., et al.: Crowdsourcing street view imagery: a comparison of mapillary and OpenStreetCam. ISPRS Int. J. Geo Inf. **9**(6), 341 (2020)
18. Hou, Y., Biljecki, F.: A comprehensive framework for evaluating the quality of street view imagery. Int. J. Appl. Earth Obs. Geoinf. **115**, 103094 (2022)
19. Fan, Z., Feng, C.-C., Biljecki, F.: Coverage and bias of street view imagery in mapping the urban environment. Comput. Environ. Urban. Syst. **117**, 102253 (2025)